

# SEARCH REQUEST FORM

Requestor's

Name: \_\_\_\_\_

Serial

Number: \_\_\_\_\_

Date: \_\_\_\_\_

Phone: \_\_\_\_\_

Art Unit: \_\_\_\_\_

## Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

## STAFF USE ONLY

Date completed: 07-29-03

Searcher: Beverly 24594

Terminal time: 22

Elapsed time: \_\_\_\_\_

CPU time: \_\_\_\_\_

Total time: 23

Number of Searches: \_\_\_\_\_

Number of Databases: 1

### Search Site

\_\_\_\_\_ STIC

\_\_\_\_\_ CM-1

\_\_\_\_\_ Pre-S

### Type of Search

\_\_\_\_\_ N.A. Sequence

\_\_\_\_\_ A.A. Sequence

\_\_\_\_\_ Structure

\_\_\_\_\_ Bibliographic

### Vendors

\_\_\_\_\_ IG

\_\_\_\_\_ STN

\_\_\_\_\_ Dialog

\_\_\_\_\_ APS

\_\_\_\_\_ Geninfo

\_\_\_\_\_ SDC

\_\_\_\_\_ DARC/Questel

☒ Other CGN



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 99905**

**To: Sarvamangala Devi**

**Location: CM1-7E12**

**Art Unit: 1645**

**Tuesday, July 29, 2003**

**Case Serial Number: 10/060521**

**From: Beverly Shears**

**Location: Biotech-Chem Library**

**CM1-1E05**

**Phone: 308-4994**

**beverly.shears@uspto.gov**

### **Search Notes**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 28, 2003, 15:35:01 ; Search time 98 Seconds

(without alignments)  
806,590 Million cell updates/sec

Title: US-10-060-521-2

Perfect score: 2540  
Sequence: 1 MSKELKQLSPAYDNQEVLL.....IEHDHFMKKITDKIVLKS 498

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

al number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 75%  
Maximum Match 100%  
Listing first 1000 summaries

Database :  
1: A\_Geneseq\_19jun03:\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2540	100.0	498	23	ABG32864 E. faecalis hygrom
2	2532	99.3	498	22	ABB47285 Enterococcus faeca

#### ALIGNMENTS

RESULT 1  
ABG32864

ID	ABG32864 standard; Protein; 498 AA.
XX	
AC	ABG32864;
XX	
DT	20-NOV-2002 (first entry)
XX	
DE	E. faecalis hygromycin A resistance protein, harA.
XX	
KW	HarA; enzyme; hygroscopic A resistance; bacterial infection; sepsis;
KW	septic shock; bacterial endocarditis; prostatitis; NTPase;
KW	staphylococcal infection; pneumococcal infection.
XX	
OS	Enterococcus faecalis strain OG1X.
XX	
PN	EP1227324-A2.
XX	
XX	31-JUL-2002.
PD	
XX	29-JAN-2002; 2002EP-0250612.
PF	
XX	30-JAN-2001; 2001US-265034P.
PR	
XX	(PF12) PFIZER PROD INC.
PA	
PI	Batma ET, Mueller JP,
XX	
XX	WPI, 2002-676579/73.
DR	N-PSDB; ABS53679.
DR	
XX	
PT	Novel hygroscopic A resistance polypeptide derived from Enterococcus
PT	faecalis and Bacillus subtilis useful for screening compounds which
PT	binds to the polypeptide, and as medicament for treating drug resistant
PT	infection -
XX	
XX	Claim 12; Fig 2; 43pp; English.
PS	
XX	The invention relates to a recombinant hygroscopic A resistance (harA)
CC	polypeptide comprising a sequence which is at least 75% identical to
CC	the harA protein form Enterococcus faecalis and Bacillus subtilis.
CC	Also included are: (1) use of a hygroscopic A-resistant strain of
CC	Enterococcus faecalis or Bacillus subtilis to determine whether an
CC	antibacterial agent is effective in treating organisms which exhibit
CC	harA-mediated drug resistance; (2) the nucleic acid encoding the harA
CC	polypeptide and (3) harA expression vectors and host cells.
CC	HarA protein or nucleic acid is useful in the screening/manufacture of
CC	a compound for use as a medicament for the treatment of a drug resistant
CC	infection. HarA nucleic acid is useful to identify an organism containing
CC	a harA gene. HarA protein is useful to screen for E. faecalis antibodies
CC	in blood or serum and samples from animals, or as antigens to raise
CC	antibodies, which antibodies are useful as diagnostic reagents. The
CC	identified compounds are useful in the treatment of enterococcal
CC	infections (e.g. sepsis, septic shock, bacterial endocarditis and
CC	prostatitis) as well as staphylococcal and pneumococcal infections.
CC	The nucleic acid is useful as a hybridisation probe for
CC	RNA, cDNA and genomic DNA to isolate full-length cDNAs and genomic
CC	clones encoding harA polypeptides and to isolate cDNA and genomic
CC	clones of other genes that have a high sequence similarity to harA
CC	gene. HarA has NTPase activity. The present sequence is the E. faecalis
CC	harA protein.
XX	
XX	
SO	Sequence 498 AA;
Query Match	100.0%; Score 2540; DB 23; Length 498;
Best Local Similarity	100.0%; Pred. No. 2.1e-190;
Matches 498; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MSKELKQLSPAYDNQEVLLPDQANTTMDTNTWGLIGRGKRTTLRLLOQLDYQGR 60
DB	1 MSKELKQLSPAYDNQEVLLPDQANTTMDTNTWGLIGRGKRTTLRLLOQLDYQGR 60
QY	ILHQVDVFFPQTVABEQQLTYVLOEYTSFQWELERELTLNVPVLMRPFSSLSGG 120
DB	ILHQVDVFFPQTVABEQQLTYVLOEYTSFQWELERELTLNVPVLMRPFSSLSGG 120

Qy	121	KTKVTLGLILEIBENAPFLIBEPHTNLDLAGQOVALYKCKKHGFIIVSHBPAFDEVY	180
Db	121	EKIVLJGLIBENAPFLIBEPHTNLDLAGQOVALYKCKKHGFIIVSHBPAFDEVY	180
Qy	181	DHILAIEKSOULTLYOGNFSIYEBOKKLRDAFELAENEKIKKDEVNRLKETARKAEWSMNR	240
Db	181	DHILAIEKSOULTLYOGNFSIYEBOKKLRDAFELAENEKIKKDEVNRLKETARKAEWSMNR	240
Qy	241	EGDKYGNKKEGSGALPRTGAI GAAPAPRWKSKHIOQRAETOLAKKYLKDLLEYIDPL	300
Db	241	EGDKYGNKKEGSGALPRTGAI GAAPAPRWKSKHIOQRAETOLAKKYLKDLLEYIDPL	300
Qy	301	SMDOYPTHHKTLITYEBRLIAGEYKMWLPAPLPSI INAGRIVGIYTKNGSGSKSLIOYLLD	360
Db	301	SMDOYPTHHKTLITYEBRLIAGEYKMWLPAPLPSI INAGRIVGIYTKNGSGSKSLIOYLLD	360
Qy	361	NPSGDSGBEATLALHOLLISYVRQDYEDNOGTLSEBPAKQOLDYTOPLANLRLKJAMERAVF	420
Db	361	NPSGDSGBEATLALHOLLISYVRQDYEDNOGTLSEBPAKQOLDYTOPLANLRLKJAMERAVF	420
Qy	421	TNRIFQMSGQKRYEVAKSLSOSAEIYIYDEPLANTLDPVNHQOLEALLISYVPAMLYIE	480
Db	421	TNRIFQMSGQKRYEVAKSLSOSAEIYIYDEPLANTLDPVNHQOLEALLISYVPAMLYIE	480
Qy	481	HDHAFMKITLDDKIVLKS	498
Db	481	HDHAFMKITLDDKIVLKS	498

RESULT 2  
ABB47285  
ID ABB47285 standard; Protein; 498 AA

AC ABB47285;

DT 31-JAN-2002 (first entry)

DE Enterococcus faecalis polypeptide Abc23.

KW MDR; efflux pump; multidrug resistance; antibacterial; drug target.

**OS Enterococcus faecalis.**

PN W0200179257-A2.

PD 25-OCT-2001.

12-APR-2001; 2001WO-US12230

PR 14-APR-2000; 2000US-197349P

PA (PHYT-) PHYTERA INC.

PI Davis DV, Rogers BL, White AC;

DR WPI; 2001-626526/72.

XX  
XX  
Detailed Report

PT	encodes/functions as
PT	encodes/functions as

PT high identity to kno  
WV

PS Claim 10; Fig 26; 139pp; English.  
XX

CC The invention relates  
CC (ABA82938-ABA93071)  
CC

CC as a multidrug resistance  
CC database for common

CC deleting/mutating a  
CC determining whether

sensitivity to an a

CC further identifying MDR efflux pumps that may be used as drug targets to  
CC increase the sensitivity of cells to antibacterial agents. Cells  
CC comprising the identified pumps may be used to screen for potential  
CC blockers or inhibitors of MDR pump function or gene expression.

**SQ. Sequence 498 AA;**

Query Match	99.3%	Score 2522;	DB 22;	Length 498;
Post 1001 54m17444	99.3%	Pred NO	5 5e-189;	

Matches	494;	Conservative	2;	Mismatches	2;	Indels	0;	Gaps	0.
---------	------	--------------	----	------------	----	--------	----	------	----

1 MSKIELKOLSPAYDNOEVLFPDANITMDTNWKGGLIGRNGRGKTLERLQKQLDYQGB 60

Db 1 MSKIELKQLSFAYDNGEALLFDQANITMDTNMKGLIGRNGRGKTIILRLLLQKQULDIYQGB 60

61 ILHQVDFVYFPQTVAABEQOLTYTVLQSVTSFEQMBLERELTLLNVDPBVLWRFSSLSGG 120

Db 61 ILHQVDFVYFPQTVAEBQQLTYTVLQEVTSFBQWKLBRRLTLNLNDPBEVLWMPFSSLSGG 120

121 EKYKVLGLLFIBNAPPLIDBPTNHLDIAGROQVAEYLGKQCKGFLVSHDRAFDVV 180

Db 121 EKT~~KL~~LG~~LI~~PIE~~NA~~P~~PL~~ID~~PT~~NH~~LD~~LAG~~Q~~V~~AB~~YLK~~Q~~Q~~Q~~G~~FI~~LVS~~HD~~RA~~PV~~DBV 180

181 DHILAEKSQTLYQGNPSIYBEQKLRDAFELAEENKIKGQVNRLLKETARKKAEWSMNR 240

Db 181 DHILAIEKSQTLTYQGNFSIYBEQKLRDAFBLAENKIKKEVNRLKETARKKAEMSMNR 240

QY 241 EGDYGNABKSGAI FDTGAI GAPAARVMKRSKHIQORAEQLAEKSKLLKDLBYIDPL 300

Db 241 EGDKYGNKKGSGAIFDTGAIGAPARVMRSKHIQQRAETQLAEKIKLLKDLBYIDS 300

301 SMDYQPTHKTLTVEBLRGYENMLPAPLSINAGBIVGITGKNSSGKSLIYL LLD 360

Db 301 SMDYQPTHTKTLTVEELRLGYEKNNLPAPISPSINAGBIVGITGKNSSGKSSLIQYLLMD 360

361 NPSGDEGEATLAHQLTISYRQDYEDNOCTLSBPAKNOUDYTOPLNNLRKLGMBRAVF 420

Db 361 NFGDSEGEATLAHQLTISVRQDYEDNQGTLSEFABKNOQLDYTOFLNNLRKLEMERAVE 420

421 TNRIBQMSMGORKVEAKSLSQSABLYIMDEPLNYLDVFNHQOLKALILSVKPAMLVIB 460

Db 421 TNRIEQMSMGQRKKVEAKSLSQSABLYIMDRPLNYLDVFNHQQLKALILS VLPATLYLB 200

QY	481	HDAHFMIKCTDICKVLKS	498

Dd 481 HDAHFMGGTIDKKVLKS 498

Search completed: July 28, 2003, 15:39:55  
Job time : 99 secs

Search completed: July 28, 2003, 15:39:55  
Job time : 99 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: July 28, 2003, 15:40:06 / Search time 30 Seconds  
(without alignments)  
702.360 Million cell updates/sec

Title: US-10-060-521-2

Perfect score: 2540  
Sequence: 1 MSKIRLQKLSFAYDNQEVLL.....IEHDAIFMKKITYDKKIVLKS 498

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

al number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 75%  
Maximum Match 100%  
Listing first 1000 summaries

## Database :

Issued Patents\_AA:\*

- 1: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*
- 2: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*
- 3: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*
- 4: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*
- 5: /cgn2\_6/prodata/1/1aa/ECTUS\_COMB.pep:\*
- 6: /cgn2\_6/prodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
------------	-------------	-------	--------	----	-------------

0 matches found

Search completed: July 28, 2003, 15:50:53  
Job time : 30 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 28, 2003, 15:38:06 ; Search time 53 Seconds  
(without alignments)  
115.894 Million cell updates/sec

Title: US-10-060-521-2

Perfect score: 2540  
Sequence: 1 MSKIEKLKLSFAYDNQEVTL.....IEHDAHFMKKTIDKKIVLKS 498

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

al number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 75%

Maximum Match 100%

Listing first 100 summaries

Database :

Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query				
No.	Score	Match	Length	ID	Description
No matches found					

Search completed: July 28, 2003, 15:48:26  
Job time : 53 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 28, 2003, 15:36:25 ; Search time 52 Seconds  
(without alignments)  
921.000 Million cell updates/sec

Title: US-10-060-521-2

Perfect score: 2540  
Sequence: 1 MSKIELKQLSPAYDNQEVLT.....IEHDAHFPMKITTDKKIVLKS 498

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

al number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 75%  
Maximum Match 100%  
Listing first 1000 summaries

Database : PIR\_76: +  
1: p1r1: +  
2: p1r2: +  
3: p1r3: +  
4: p1r4: +

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
---------------	-------	----------------	--------	----	-------------

No matches found

Search completed: July 28, 2003, 15:40:58  
Time : 52 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 28, 2003, 15:35:01 ; Search time 25 Seconds  
(without alignments)  
936.772 Million cell updates/sec

Title: US-10-060-521-2

Perfect score: 2540

Sequence: 1 MSKILKQLSPAYDNOEVL.....IEHDAHPMKITDKKIVLKS 498

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

al number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 75%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Score	Match	Length	ID	Description
---------------	----------------	-------	--------	----	-------------

No matches found

Search completed: July 28, 2003, 15:38:21  
Job time : 33 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 28, 2003, 15:38:26 ; Search time 103 Seconds  
(without alignments)  
1247.672 Million cell updates/sec

Title: US-10-060-521-2

Perfect score: 2540  
Sequence: 1 MSKRLKQLSFAYDQEVLL.....IEHDAHFMKKITDKKIVLKS 498

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

al number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 75%  
Maximum Match 100%  
Listing first 1000 summaries

Database :

SPTREMBL\_23: \*  
1: sp\_archaea: \*  
2: sp\_bacteria: \*  
3: sp\_fungi: \*  
4: sp\_human: \*  
5: sp\_invertebrate: \*  
6: sp\_mammal: \*  
7: sp\_mhc: \*  
8: sp\_organelle: \*  
9: sp\_phage: \*  
10: sp\_plant: \*  
11: sp\_rodent: \*  
12: sp\_virus: \*  
13: sp\_vertebrate: \*  
14: sp\_unclassified: \*  
15: sp\_virus: \*  
16: sp\_bacteriap: \*  
17: sp\_archaeap: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
No matches found					

Search completed: July 28, 2003, 15:50:22  
Job time : 110 secs